

## **EXHIBIT I**

## Wendy Thai

---

**Subject:** FW: [help #21742] RE: GROUP:Q9H5V8  
**Attachments:** Wendy Thai.vcf

-----Original Message-----

From: Elisabeth Gasteiger via RT [<mailto:help@uniprot.org>]  
Sent: Wednesday, June 20, 2007 4:15 AM  
To: Wendy Thai  
Subject: [help #21742] RE: GROUP:Q9H5V8

> [WThai@slwk.com - Wed Jun 20 00:25:32 2007]:  
>  
> Hi Elisabeth:  
>  
> Thank you for your help with this.  
>  
> I was told by an examiner at the US Patent and Trademark Office that the  
> Scherl-Mostageer sequence had been corrected.  
>  
> Also, my sequence analysis data (see attached WORD document) indicates  
> that the Scherl-Mostageer sequence has been changed in the database.  
>  
> For example, when I do an alignment of the original Scherl-Mostageer  
> sequence with my sequence (SIMA135), I see two amino acid mismatches -  
> at positions 525 and 827. See page 16 of the attached WORD document.  
>  
> However, when I do a blastp analysis of my sequence (SIMA135) against  
> the database, I get 100 % identity with Q9H5V8. See page 1 of the  
> attached WORD document. Q9H5V8 references the Scherl-Mostageer sequence  
> (reference 1) - SEE pages 7-8 of attached WORD document.  
>  
> Thus, it appears as if the Scherl-Mostageer sequence had been changed  
> and is now identical to SIMA135. Please advise. Am I misreading the  
> sequence analysis results? If you think a telephone call would be  
> helpful, let me know and I can call you.  
>  
> Thanks so much for your help.  
>  
> Wendy Thai  
> 612-373-6913

Dear Wendy,  
Thank you for these precisions.  
The 2 mismatches are indeed reported in the UniProtKB/Swiss-Prot record.  
The first one is a known variant:  
FT VARIANT 525 525 R -> Q (in dbSNP:rs3749191).  
FT /FTId=VAR\_025498.

And the second is annotated as a sequencing conflict (ref.1 is the  
Scherl-Mostageer paper in the entry):  
FT CONFLICT 827 827 S -> N (in Ref. 1, 3 and 5; BAB14695).

Your confusion is indeed caused by the fact that **Swiss-Prot is a non-redundant database that strives to have 1 protein entry for each gene product. When the sequence was manually annotated, the scientist who did the annotation saw that Q9H5V8 and Q96QU7 (the TrEMBL entry translated from the Scherl-Mostageer sequence AY026461) described the same protein although their sequences differed slightly (in 2 residues).**

And here is where the answer I gave you yesterday was not quite correct:  
I said the sequences were 100% identical but indeed they are not.  
Here is the entry history  
<http://www.ebi.ac.uk/uniprot/unisave/?query=Q9H5V8&search=Go>

and here is the portion of our user manual which describes the concept  
of minimal redundancy:  
[http://www.expasy.org/sprot/userman.html#what\\_is\\_sprot](http://www.expasy.org/sprot/userman.html#what_is_sprot)

In summary, the Scherl-Mostageer sequence has not changed since its  
submission. It is just in the manually annotated Swiss-Prot database  
that it has been considered to describe the same protein as the TrEMBL  
sequence with which it was merged (note that this is an old version of  
the Q9H5V8 entry from before the merge - it was in TrEMBL then and moved  
to Swiss-Prot upon manual annotation):

ID Q9H5V8\_HUMAN PRELIMINARY; PRT; 836 AA.  
AC Q9H5V8;  
DT 01-MAR-2001, integrated into UniProtKB/TrEMBL.  
DT 01-MAR-2001, sequence version 1.  
DT 07-FEB-2006, entry version 13.  
DE Hypothetical protein FLJ22969 (NCSG135).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,  
RA Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T.,  
RA Nakamura Y., Isogai T., Sugano S.;  
RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=22547370; PubMed=12660814; DOI=10.1038/sj.onc.1206220;  
RA Hooper J.D., Zijlstra A., Aimes R.T., Liang H., Claassen G.F.,  
RA Tarin D., Testa J.E., Quigley J.P.;  
RT "Subtractive immunization using highly metastatic human tumor cells  
RT identifies SIMA135/CDCP1, a 135 kDa cell surface phosphorylated  
RT glycoprotein antigen.";  
RL Oncogene 22:1783-1794(2003).  
CC -----  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC -----  
DR EMBL; AK026622; BAB15511.1; -; mRNA.  
DR EMBL; AF468010; AAO33397.1; -; mRNA.  
DR Ensembl; ENSG00000163814; Homo sapiens.  
SQ SEQUENCE 836 AA; 92875 MW; 9B980475C3E5C4C8 CRC64;  
MAGLNCGVSI ALLGVLLGA ARLPRGAEAF EIALPRESNI TVLIKLGTP TLLAKPCYIVI  
SKRHITMSLS KSGERIVFTF SCQSPENHFV IEIQKNIDCM SGPCPFGEVQ LQPSTSL LPT  
LNRTFIWDVK AHKSIGLELQ FSIPRLRQIG PGESCPDGV T HSISGRIDAT VVRIGTFCSN  
GTVSRIKMQE GVKMALHLPW FHPRNVSGFS IANRSSIKRL CIIESVFEGE GSATLMSANY  
PEGFPEDELM TWQFVPAHL RASVSFLNFN LSNCKERKEER VEYYPGSTT NPEVFKLEDK  
QPGNMAGNFN LSLQGCDQDA QSPGILRLQF QVLVQHPQNE SNKIYVVDLS NERAMSLTIE  
PRPVKQSRKF VPGCFVCLES RTCSSNLT LTSGSKHKISFL CDDLTRLWMN VEKTISCTDH  
RYCQRKSYSL QVPSDILHLP VELHDFSWKL LVPKDRLSLV LVPAQKLQQH THEKPCNTSF  
SYLVASAIPS QDLYFGSFCP GGSIKQIQVK QNISVTLRTF APSFRQEASR QGLTVSFIPI  
FKEEGVFTVT PDKSKVYLR TPNWDRGLPS LTSVSWNISV PRDQVACLTF FKERSGVVCQ  
TGRAFMIIQE QRTRAEIIFS LDEDVLPKPS FHHHSFWVNI SNCSPSTSGKQ LDLLFSVTLT  
PRTVDLTVIL IAAVGGGVLL LSALGLIICC VKKKKKKTKN GPAVG IYNGN INTEMPRQPK  
KFQKGRKDND SHVYAVIEDT MUYGHLLQDS SGSFLQPEVD TYRPFQGTMG VCPPSPPTIC  
SRAPTAKLAT EEP PPRSPPE SESEPYTF SH PNNGDVSSKD TDIPLLSTQE PMEPAE

//

I hope this helps.

Best regards  
Elisabeth Gasteiger

--

---

Elisabeth Gasteiger  
Swiss Institute of Bioinformatics  
CMU - 1, rue Michel Servet  
CH - 1211 Geneva 4 Switzerland  
Elisabeth.Gasteiger@isb-sib.ch

Tel. (+41 22) 379 58 79  
Fax (+41 22) 379 58 58  
<http://www.expasy.org/>



Wendy Thai.vcf  
(182 B)

---